We thank Jorge Ramirez for his review of our manuscript and making the highly constructive comments and suggestions. We are glad to hear our effort to revise the manuscript. The comments and suggestions were formatted in light and dark blue text. The author's response to the major issues is shown below in **black text**.

General comments

In this study a landscape evolution model, CAESAR-Lisflood (CL), is applied to a steep mountain catchment to assess the effectiveness of engineering works in reducing the transport of sediment. This is an applied study, that is straightforward, and demonstrates the use of CL in a highly dynamic landscape. Overall, the manuscript fits the scope of NHESS and would be interesting to modellers and practitioners working in mitigating geo hazards in mountainous regions. My concerns with the study are related to the choice of hydrological parameters, the physical plausibility of landscape changes, and the development of initial conditions. In addition, the clarity of the manuscript requires substantial improvement and I recommend the text is thoroughly edited by a native English speaker, with a background in fluvial geomorphology, before acceptance. Below are major comments that need to be addressed followed by a list of minor points and edits.

Major comments

A weakness of the study is the lack of **calibration of the hydrological component in CL**. As such, there is no way of knowing if the quantity and timing of the floods in the ungauged catchment are accurately replicated by CL. The hydrological parameters adopted (m-values) are from studies performed from nearby catchments but these studies also have not performed calibration to derive m-values. The authors, instead rely on landcover to assign m-values, but **m-value is only partly dependent on landcover**. For example, Ramirez et al.2022 found that in a mountain catchment **soil depth correlated well with m-value** and not with landcover. To have greater confidence in the model, the authors need to provide **hydrographs for the entire simulated period** and, in addition, provide qualitative or quantitative data that confirms the physical plausibility of the simulated discharge, specifically the floods.

1. *The m-values*

The m-values in C-L influence the peak and duration of the hydrograph in response to rainfall (Coulthard et al., 2002), which are usually determined by the landcover (e.g., 0.02 for the forest, 0.005 for the grassland) (Coulthard and Van De Wiel, 2017). In our study, we united the value to 0.008 in our smaller catchment (14 km^2) in Scenario UP and PP, which resembles the m-value of farmland covered with lower vegetation in the same catchment studied by Xie et al., (2022), Li et al., (2020) and Xie et al., (2018). In scenario EP, the m-value in the vegetation revetments area was 0.02 to distinguish the vegetation coverage. It has been calibrated in the bigger catchment containing our study area (Xie et al., 2018) by replicating the flood event in 2013.

We have read Ramirez et al., (2022) intensively and learned the new and creative method to calibrate components. They determined m-values in the total catchment after perfect simulation results in a sub-catchment according to the soil depth. And we think

It would exist discrepancies between different regions. In our catchment, the vegetation counts more than the soil depth for the m-value, which is caused by the undetermined soil depth. Therefore, it is not the optimal method for our smaller study area distributed larger amount of landslides deposition and river alluvium stem from metamorphic sandstones and sandy slate.

2. Calibration

Admittedly, it is not enough for our calibration work including referencing the parameters from the published research in the same catchment and using the recommended values by model developers. Considering the ungauged basins before 2015, we replicated the flash flood event in July 2018 by C-L to calibrate the hydrological components.

There are no huge differences in geomorphology, channel location, and landcover before 2013 and after 2018 in our catchment found from the field surveys. Based on Scenario PP (with two check dams), we changed the rainfall series to the two-week hourly precipitation in July 2018, which is recorded by the rain gauge 2.5 km away from the catchment placed in 2015. The simulation results (Figure 1c and Figure 1d) showed the erosion and maximum flood depth deposition distributions in Scenario PP on July 15th, 2018. As shown in Figure 1c and Figure 1d, we selected three locations randomly to compare the simulation results with remotely sensed images and photos. The results (Figure 2) showed reliable results including sediment deposition and the peak flood depth, which indicate that the flash flood event was replicated successfully by the C-L.

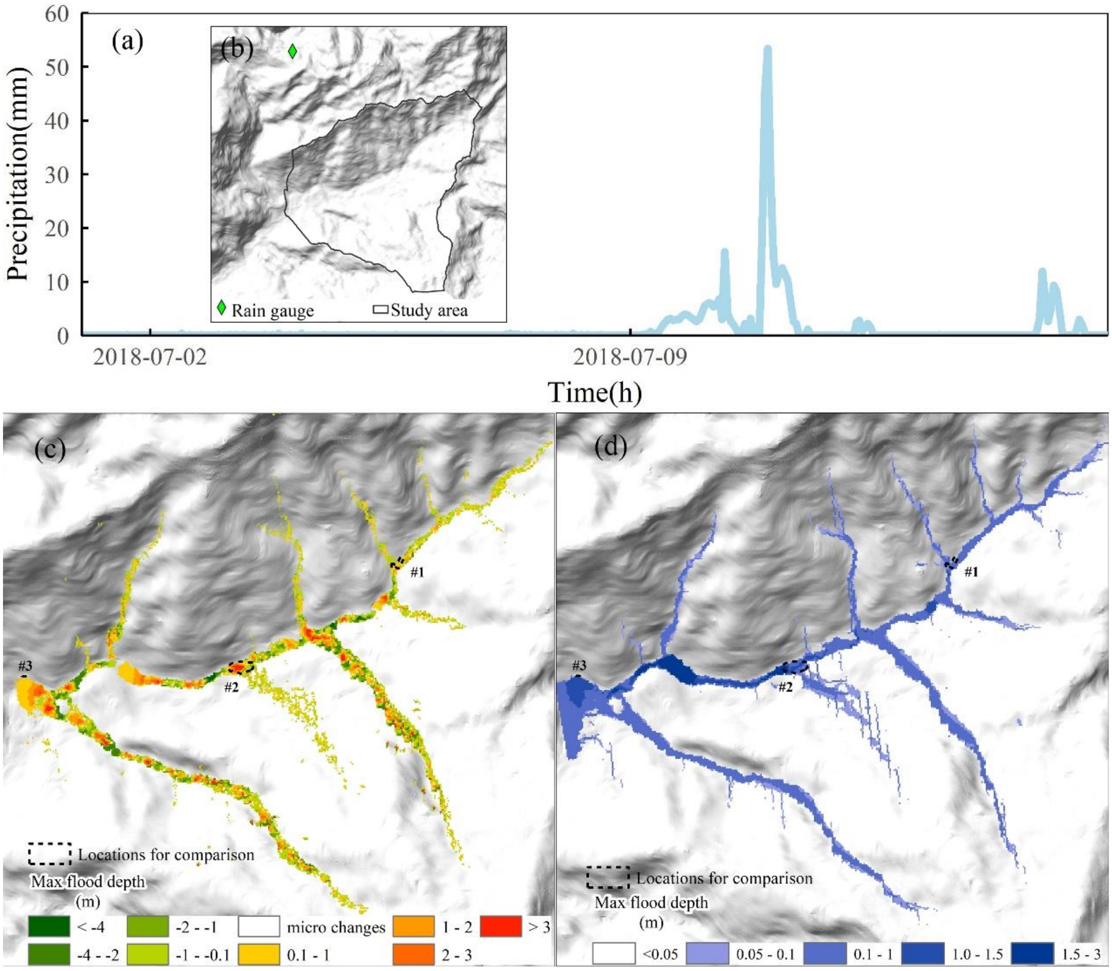

Figure 1. The input rainfall series (a and b) and simulation results of the flash flood event in July 2018 (c and d).

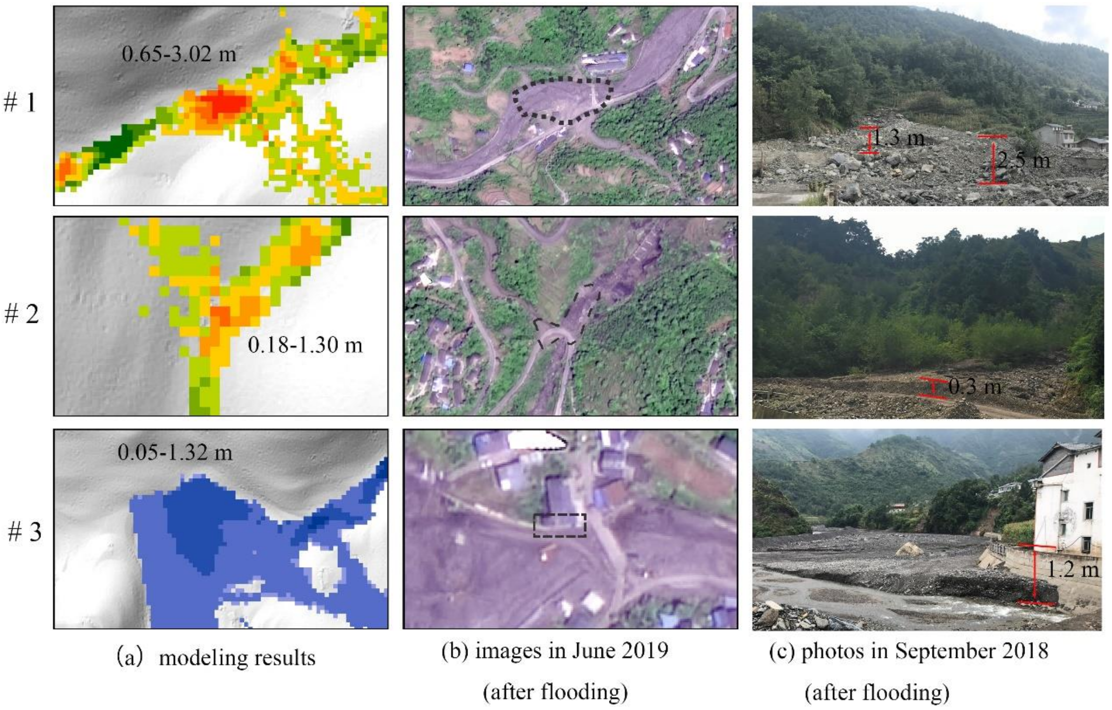

Figure 2. The comparison of the simulation results to images (GF-2 with 8-m resolution) and photos after the flash flood event in July 2018.

3. Verification

In section 5.1, we confirmed the plausibility of the simulated results using deposition depth evaluated from photos. Herein, we add the discharge of the entire simulation period for Scenario PP. As shown in Figure 3, we compare two types of discharge recorded in published research (Feng et al., 2017; Guo et al., 2018) with those of simulation results to confirm the physical plausibility. And we captured five flood events where the daily precipitation is more than 50 mm in 2013 and the peak discharge was up to 63.6 m³/s.

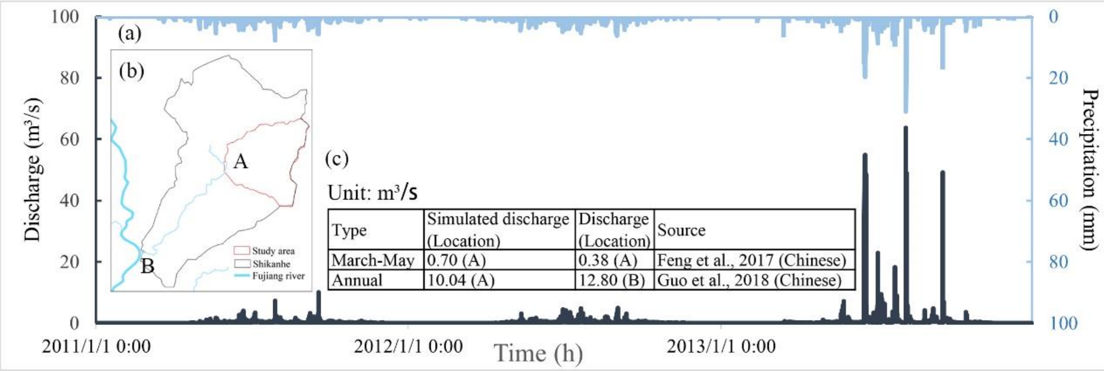

Figure 3. (a) The simulation discharge in 2011-2013 in Scenario PP; (b) the verification location; (c) the comparison of the simulated to the recorded discharge.

In this study, CL simulations have produced locations of deep erosion between 3-10 m in a period of three years. This is quite a bit of erosion in such a short period and in some instances would produce features in the simulated landscape that resemble small canyons. Could you **verify that these erosional features are physically plausible** by providing photographic evidence from the observed landscape and comparing them to cross-sections from the simulation? Or provide any other type of validation that supports such extreme erosion across the simulated landscape. In addition, across all simulations (Fig. 5a), there are instances of erosion that exceed 3 m in the downstream area where erodible thickness is 3 m. Can the authors explain how **simulated erosion can exceed the thickness of the initial erodible sediment?** Likewise, in this study, how is it possible for CL to produce erosion between 10 and 15 m, if the maximum depth of erodible sediment in the catchment is 10 m?

4. Revised erosional and deposited features

Many thanks for your reminding and we have to apologize for our mistakes to show the abandoned simulation result, where the input basedDEMs were generated improperly. Now we update the revised figures as shown in Figure 4 and Figure 5, which both show the spatial distribution of erosion and deposition. We correct the description (extreme erosion (<-7 m), heavy erosion (-7--3 m), moderate erosion (-3--1 m), light erosion (-1-0.1 m), micro change (-0.1-0.1 m), light deposition (0.1-1 m), moderate deposition (1-3 m), heavy deposition (3-7 m), and extreme deposition (>7 m)) in section 4.1 and 4.2. We ensure that all the analysis results and input parameters are consistent and from our optimal simulation after checking all the figures, tables and numbers.

In addition, we thank the reviewer for his suggestions about the figure's details.

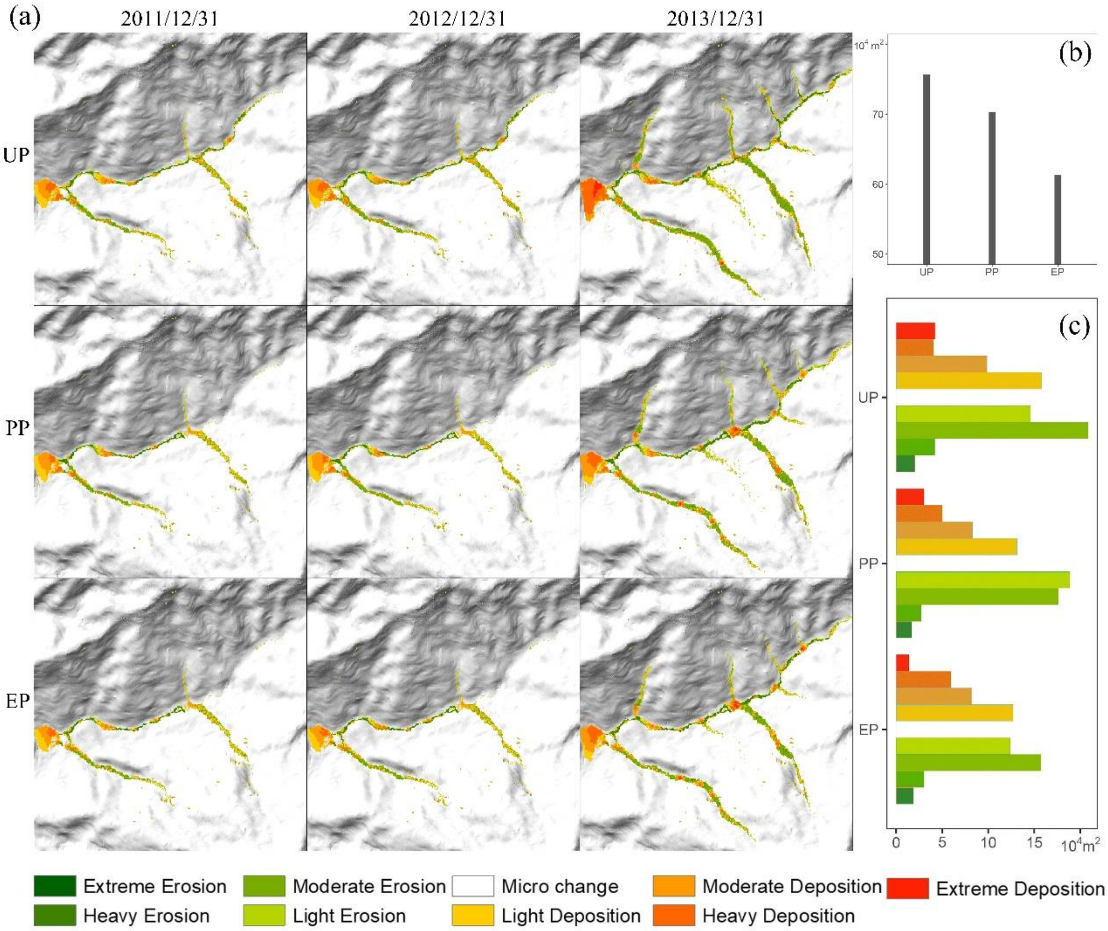

Figure 4. (a) Simulated geomorphic changes over time for three scenarios; (b) the exposure area included deposition and erosion for three scenarios; (c) the distribution of deposition and erosion at the conclusion of the simulation for the three scenarios.

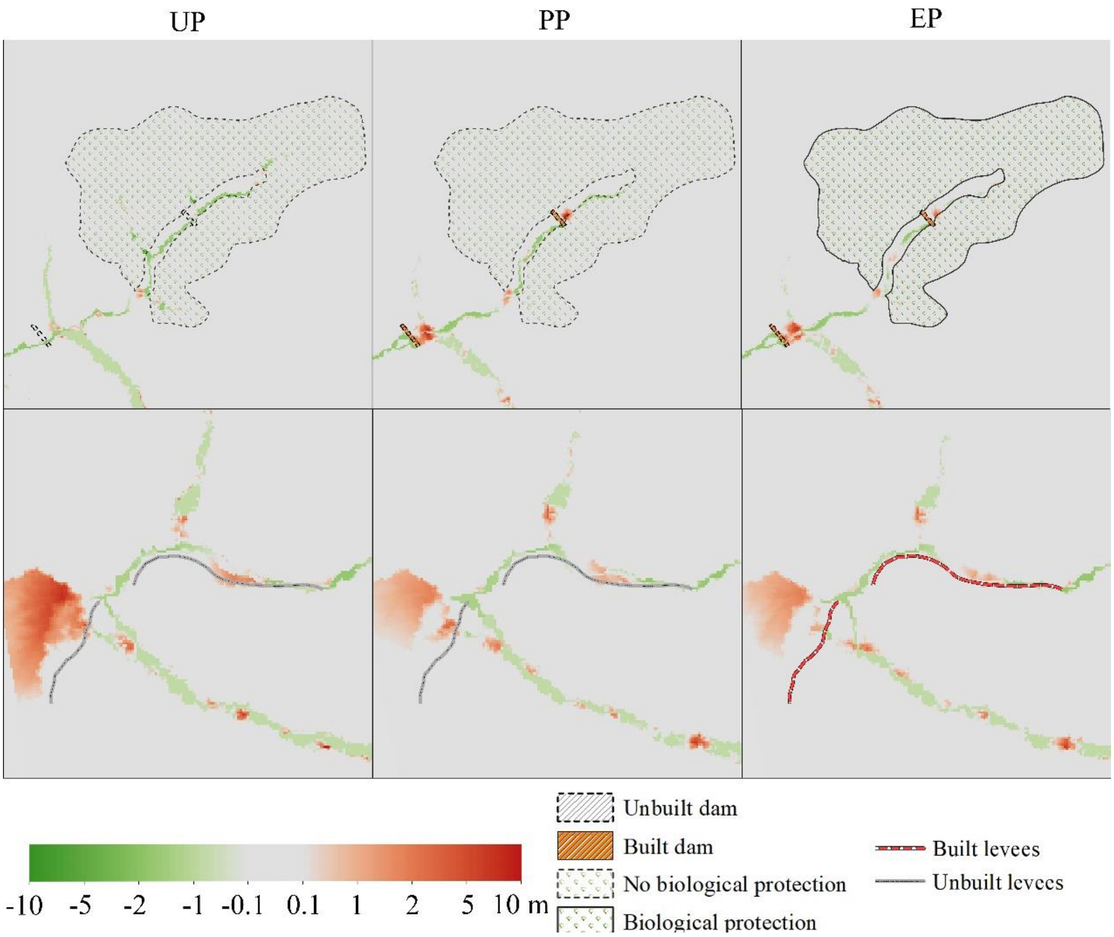

Figure 5. Geomorphic changes at the conclusion of the simulation at key spots for the UP, PP, and EP scenarios. Top row is the upriver section containing dam 1, dam 2 and the vegetation revetment. The bottom row is the downriver section containing levees.

5. *Verification of erosion and deposition*

As shown in Figure 6, we verify the erosional and deposited features by providing photographic evidence from the observed landscape and compare them to cross-sections from the simulated results.

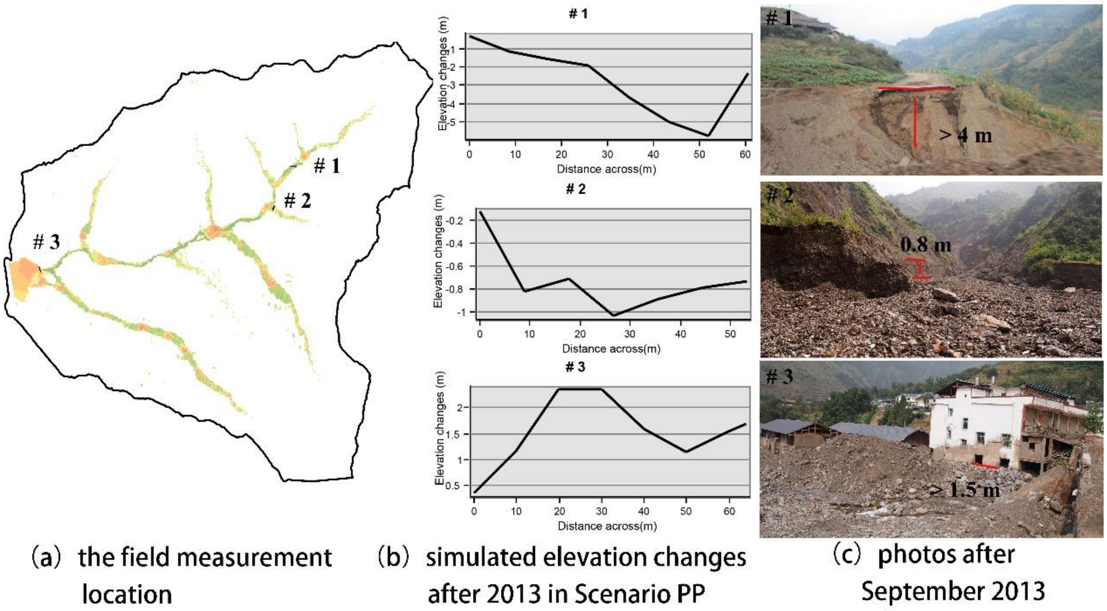

Figure 6. The comparison of cross-sections from the simulation results to the photos in the field measurement locations after 2013 in Scenario PP.

In the study, there is no mention of establishing initial conditions by spinning-up the model to mix the grain sizes. **If spin-up was not performed, can the authors provide an explanation.** If spin-up was performed, could you briefly explain how it was done in the methods. Regarding the choice of bedrock elevation (Fig. 2), could the authors provide the physical basis for the choice of erodible thickness values and locations of these values.

6. *Spin-up processes*

Admittedly, we didn't spin up the model to mix the grain sizes. The purposes of the process are to eliminate the 'walls' and the 'depressions' in the cells and avoid the intense erosion in the hill slope in the early run time. Actually, we preprocess the DEMs by filling sink based on Environmental Systems Research Institute's (ESRI's) ArcMap (ArcGIS, 10.8) to eliminate the problematic pixels. Moreover, for our catchment, the fine grains distributed homogeneously both in the hill slope and the channel five years after the strong earthquake. Therefore, we think the huge difference would not exist. However, we will continue to compare the difference in the future work.

7. *Erodible thickness values*

The bedrock elevation (Fig. 2) was evaluated mainly from the published research from the trusted official institutions in China (Feng et al., 2017; Guo et al., 2018) and verified with our field survey. The institutions described in their research according to the prompt and accurate hazard inventory and the UAV survey. It is difficult for us to provide direct evidence like drilling operations and the fine map because of the steep terrain and a large amount of landslides deposition in these post-earthquake fragile mountains.

Minor points and edits

The most minor points and edits would be revised in the manuscript directly and some are reply here.

Line 145-146: Provide an example of a vegetation revetment.

Considering the tree roots play an important role in stabilizing the slope and consolidating the soil, the ecological engineering including vegetation revetment was more and more popular in the mountains. For example, the tree planting patterns was studied by Lan et al., (2020). They listed the artificially planted cypress and pines on the slope.

Line 269-270: Here you mention "damage" but what you are really calculating is exposure because you are not calculating a monetary value. Change the text to mention exposure, and I am assuming that a map of settlements or landcover was used to calculate the exposure, if so, provide this map in the supplemental information. If you didn't use a landcover map, explain what you used to derive exposure.

Actually, we only calculate the total of deposited and erodible area in each scenario at the conclusion of the simulation to compare.

In Figure 6, the EP map shows the levees blocking a tributary, is this a mistake in the figure or did you really block this tributary in the simulation. Please explain.

We have checked carefully and ensured that the deep and narrow outlet of the tributary was not blocked by the levees.

Conclusion section: Here you need to further summarize your main findings because the current text reads like a repetition of the results.

Thanks for your suggestion. Admittedly, our conclusion is verbose and we would summarize the conclusion in the revised manuscript.

Reference

Coulthard, T. J. and Van De Wiel, M. J.: Modelling long term basin scale sediment connectivity, driven by spatial land use changes, Geomorphology, 277, 265–281, https://doi.org/10.1016/j.geomorph.2016.05.027, 2017.

Coulthard, T. J., Macklin, M. G., and Kirkby, M. J.: A cellular model of Holocene upland river basin and alluvial fan evolution, Earth Surf. Process. Landforms, 27, 269–288, https://doi.org/10.1002/esp.318, 2002.

Feng, W., He, S., Liu, Z., Yi, X., and Bai, H.: Features of Debris Flows and Their Engineering Control Effects at Xinping Gully of Pingwu County, J. Eng. Geol., 25, 2017.

Guo, Q., Xiao, J., and Guan, X.: The characteristics of debris flow activities and its optimal timing for the control in Shikan River Basin Pingwu Country, Chinese J. Geol. Hazard Control, 29, 2018.

Lan, H., Wang, D., He, S., Fang, Y., Chen, W., Zhao, P., and Qi, Y.: Experimental study on the effects of tree planting on slope stability, Landslides, 17, 1021–1035, https://doi.org/10.1007/s10346-020-01348-z, 2020.

Li, C., Wang, M., Liu, K., and Coulthard, T. J.: Landscape evolution of the Wenchuan earthquake-stricken area in response to future climate change, J. Hydrol., 590, 125244, https://doi.org/10.1016/j.jhydrol.2020.125244, 2020.

Ramirez, J. A., Mertin, M., Peleg, N., Horton, P., Skinner, C., Zimmermann, M., and Keiler, M.: Modelling the long-term geomorphic response to check dam failures in an alpine channel with CAESAR-Lisflood, Int. J. Sediment Res., 37, 687–700, https://doi.org/10.1016/j.ijsrc.2022.04.005, 2022.

Xie, J., Wang, M., Liu, K., and Coulthard, T. J.: Modeling sediment movement and channel response to rainfall variability after a major earthquake, Geomorphology, 320, 18–32, https://doi.org/10.1016/j.geomorph.2018.07.022, 2018.

Xie, J., Coulthard, T. J., and McLelland, S. J.: Modelling the impact of seismic triggered landslide location on basin sediment yield, dynamics and connectivity, Geomorphology, 398, 108029, https://doi.org/10.1016/j.geomorph.2021.108029, 2022.